

09/486799
PCT/EP 99/05460
BUNDESREPUBLIK DEUTSCHLAND

REC'D 13 NOV 1998

WIPO PCT



31

Bescheinigung

Die Océ Printing Systems GmbH in Poing/Deutschland hat eine Patentanmeldung unter der Bezeichnung

"Verfahren und Gerät zur elektronischen Archivierung eines Computer-Datenstroms"

am 29. August 1997 beim Deutschen Patentamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patentamt vorläufig das Symbol G 06 F 17/40 der Internationalen Patentklassifikation erhalten.

München, den 6. Oktober 1998
Der Präsident des Deutschen Patentamts

Im Auftrag

enzeichen: 197 37 811.0

 Sleck

**PRIORITY
DOCUMENT**

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Verfahren und Gerät zur elektronischen Archivierung eines Computer-Datenstroms

5 Die Erfindung betrifft ein Verfahren und ein Gerät zur elektronischen Archivierung des von einem Computer abgegebenen Datenstroms nach dem Oberbegriff des Patentanspruchs 1.

10 Ein solches Verfahren und ein solches Gerät sind aus der DE-4408327 A1 bekannt. In Computern erzeugte Dokumente, die sowohl grafische als auch numerische und alphanumerische Informationen enthalten, werden dabei über eine Schnittstelle, beispielsweise eine Standard-

15 Druckerschnittstelle, an ein Archivierungssystem übertragen. Innerhalb des Archivierungssystems werden die zu den Dokumenten gehörenden Daten auf Massenspeichern wie Magnetbändern, magnetischen oder optischen Speicherplatten oder dergleichen langfristig aufbewahrt. Diese Art der

20 Speicherung ersetzt in zunehmendem Maße ältere Verfahrensweisen, bei denen auf Papier vorliegende Dokumente bzw. Originale mit einem optischen Abtaster (Scanner) erfaßt, das so erhaltene Bild in elektrische Signale umgewandelt wurden und dann in Archiven abgelegt wurden. Die früher

25 übliche Microverfilmung von Dokumenten wird durch diese neue Technik in zunehmendem Maße ersetzt, wenn die Originale bereits in Form von elektronischen Signalen bzw. Computer-Datenströmen vorliegen.

30 Elektronische Archivierungssysteme der eingangs genannten Art setzen einen vom Computer ausgegebenen Datenstrom in der Regel in einen Datenstrom mit speziellem, an das Archiv angepaßten Datenformat um. In vielen Anwendungsfällen ist der

Datenstrom des Computers an bestimmte Ausgabesysteme, insbesondere an Drucker, angepaßt. Beispiele solcher Druckdatenströme sind das von IBM geprägte Format IPDS oder das von Hewlett Packard geprägte Datenformat PCL.

5

Ein Archivierungssystem setzt diese Datenströme auf ein Format um, welches dem Archivierungssystem entspricht. Die direkte Speicherung der vom Computersystem abgegebenen Daten erweist sich dabei als sehr nachteilig, weil dabei eine sehr große Anzahl an Systemparametern, z.B. Zeichenfonts, mit abgespeichert werden müßte. Auch die Reproduktion derart abgespeicherter Daten würde sich dann als sehr aufwendig erweisen. Archivierungssysteme speichern die Daten deshalb auf einer Pixel-orientierten Basis, z.B. im sog. TIF-Format, ab. Durch eine derartige, punktweise aufgebaute Speicherungstechnik eröffnet sich dann die Möglichkeit, das Datenvolumen nach gängigen Komprimierungsverfahren zu verringern. Das Datenvolumen derart komprimierter Daten ist um so größer, je mehr schwarz-weiß Übergänge zu verarbeiten sind. Andererseits werden Druckseiten in zunehmendem Maße optisch immer aufwendiger gestaltet. Beispielsweise werden auf Formularen immer häufiger Grauraster hinterlegt, um diese optisch ansprechender zu gestalten und Hervorhebungen besser kenntlich zu machen. Diese Entwicklung hat zur Folge, daß bei der Archivierung Druckseiten einen immer größeren Speicherbedarf haben.

In einem anderen, bekannten System, dem sog. COLD-System, werden Daten zur Archivierung getrennt als überwiegend grafische Daten und als überwiegend kodierte Daten (Zeilendaten) abgelegt. Bei erweiterten COLD-Verfahren werden Rohdaten und Ressourcen ebenfalls getrennt abgelegt und bei

der Reproduktion wird der gesamte Druckprozess nachgebildet.
Dies bedingt eine komplexe Ressourcen-Verwaltung.

Aufgabe der Erfindung ist es, ein System zur Archivierung von
5 Computer-Datenströmen bereit zu stellen, bei dem ein hoher
Datenkompressionsgrad erzielt werden kann, auch wenn
aufwendige grafische Informationen archiviert werden sollen.

10 Diese Aufgabe wird durch die im Patentanspruch 1 beschriebene
Erfindung gelöst. Vorteilhafte Ausführungsformen der
Erfindung sind Gegenstand der Unteransprüche.

Die Erfindung beruht auf der Erkenntnis, daß ein Großteil der
zu archivierenden Computer-Datenströme, insbesondere der an
15 Drucker zugeführten Datenströme, aus zweierlei Datenarten
besteht. Einerseits sind in diesen Druckdaten häufig
Formulare hinterlegt, die für eine bestimmte Anzahl
zusammengehöriger Druckdatensätze, sog. Jobs, in ein und
derselben Weise wiederkehren. Andererseits sind diese
20 Formulare mit variablen Daten gefüllt, welche jeweils das
Charakteristische eines Dokuments darstellen. Erfindungsgemäß
wurde erkannt, daß der durch Formulare hervorgerufene Anteil
an komprimiertem Druck-Datenstrom oft denjenigen Anteil
überwiegt, der durch die variablen Daten hervorgerufen wird.
25 Erfindungsgemäß werden deshalb diejenigen Daten, die von
Formularen stammen, von denjenigen Daten, die variabel sind,
getrennt, und die beiden Datenarten jeweils getrennt auf Bit-
Map-Basis, insbesondere komprimiert, weiterverarbeitet.
Innerhalb des Archivierungssystems werden diese Daten
30 voneinander getrennt abgespeichert und später zum Betrachten
wieder zusammengeführt; dies kann insbesondere Pixel-genau
erfolgen. Für eine Gruppe zusammengehörender Daten (Job)
werden die Daten eines Formulars nur einmal abgelegt. Bei den

einzelnen Dokumenten werden die zugehörigen variablen Daten mit dem Formular nur noch in der Weise verknüpft, daß ein Verweis auf den entsprechenden Speicherbereich des Formulars abgelegt wird. Vorteilhaft ist dabei, daß sich die - in der

5 Regel umfangreiche Datenmenge eines Formulars - für jeden Job auf ein einmaliges Abspeichern des Formulars beschränkt. Hierdurch wird eine beträchtliche Speicherplatzeinsparung gegenüber Systemen des Standes der Technik erzielt.

10 Bei Druckgut mit hoher Auflage, d.h. mit häufig wiederkehrenden, gleichartigen Formularen, wird hierdurch ein sehr hoher Grad an Speicherplatzeinsparung erreicht.

Vorteilhaft ist weiter, - insbesondere gegenüber dem o.g.

15 COLD-Systemen - daß sich die Reproduktion der Daten sehr einfach gestaltet, weil auf Bit-Map-Basis (z.B. TIF-Format) nur ein Viewer zur Reproduktion benötigt wird (selbsttragendes Format).

20 Als vorteilhaft erweist es sich auch, zusammen mit den Verweisen eine Information abzuspeichern, die ein genaues Zusammenführen der Formulardaten mit den zugehörigen, variablen Daten ermöglicht.

25 In einer weiteren, vorteilhaften Ausführungsform der Erfindung werden die Formular-Daten je Job, d.h. je zusammenhängendem Datenstrom, nur einmal im Archivspeicher als Bit Map abgespeichert, während die variablen Daten jedes Dokuments des Jobs jeweils einzeln abgespeichert werden.

30 In einer auf extrem hohen Komprimierungsgrad ausgerichteten Ausführungsform werden die Formulardaten inhaltlich reduziert oder sogar gänzlich unterdrückt. In einem derartigen System

würden dann im Archivspeicher nur noch die variablen Daten abgespeichert.

5 Im Zuge der Suche der Daten in dem Archivspeicher können die Formulardaten mit den Inhaltsdaten entweder automatisch überlagert werden anhand von Steuerungssignalen oder aber variable Daten und Formulardaten getrennt voneinander geladen und vom Bediener gesteuert übereinander gelegt werden.

10 Im Zuge des Archivierens können neben den variablen Daten und den Formulardaten sog. Index-Daten erzeugt werden, in denen Suchkriterien abgespeichert werden, sowie die Verweisdaten auf die Speicherorte der variablen Daten und gegebenenfalls der Formulardaten.

15 In einer weiteren, vorteilhaften Ausführungsform der Erfindung erfolgt der Datenübertrag zum Computersystem in das Archivsystem über einen Drucker-Controller, welcher mit einem nachgeschalteten Computer zusammenarbeitet. Vom Controller zum Computer erfolgt der Datentransfer sowohl für die Formular- und variablen Daten des punktorientierten Datenformats (Bit Map) als auch die im zeichenorientierten (ASCII-Format) Indexdaten über eine gemeinsame Schnittstelle.

20 25 Die Verknüpfung, d.h. der Verweis von den variablen Daten zu den Formulardaten, kann innerhalb des variablen Datensatzes hinterlegt werden, innerhalb des entsprechenden Index-Datensatzes oder in beiden Datensätzen.

30 Weitere Vorteile und Wirkungen der Erfindung werden anhand der nachfolgend beschriebenen Ausführungsbeispiele deutlich, die im Zusammenhang mit zwei Figuren beschrieben werden.

Es zeigen:

Figur 1: ein elektronisches Archivierungssystem und

5

Figur 2: elektronische Komponenten des Archivierungssystems.

Bei dem in Figur 1 dargestellten elektronischen Archivierungssystem wird der von einem Host-Computer 2 abgegebene Datenstrom über eine Archivierungs-Schnittstelle 1 an einen Archiv-Speicher 3 abgegeben. Der Archiv-Speicher 3 kann insbesondere ein Gerät zum Beschreiben von Laser-Speicherplatten sein. Die im Archivspeicher 3 gespeicherten Daten können in den Bit-Map-basierten Datenformat mit einem Lesegerät (Viewer) 4 abgerufen und beispielsweise wiederum auf einem Computerbildschirm dargestellt werden.

Der vom Host-Computer abgegebene Datenstrom 2 wird über eine Standard-Druckerschnittstelle in einen Controller 6 (ISTREAM) eingespielt. Diese Daten sind in Form sog. Jobs gruppiert. Ein Job ist ein zusammenhängender Datenstrom, der in irgendeiner Weise (logisch oder physikalisch) miteinander verknüpft ist. Innerhalb des Host-Computers kann beispielsweise eine Marke von außen zugeführt werden, die bestimmte Druck-Datenströme miteinander zu einem Job verbindet.

Innerhalb der Archivierungs-Schnittstelle werden die vom Host-Computer gelieferten Daten 5 klassifiziert nach variablen Daten 9 und Formular-Daten 8. Die Formular-Daten (Overlays) stellen dabei einen fixen Anteil des Jobs dar, der auf mehreren Druckseiten des Jobs unverändert auftritt. Die

variablen Daten sind dagegen innerhalb eines Jobs von Druckseite zu Druckseite unterschiedlich.

In der Archivierungs-Schnittstelle werden nunmehr für jeden 5 Job die Formulardaten 8 für jedes Formular nur jeweils einmal hinterlegt, während die variablen Daten seitenweise hinterlegt werden. Gleichzeitig werden Index-Daten erzeugt, die zu den variablen Daten bestimmte Zusatzinformation enthalten, beispielsweise herausgegriffene Ordnungsnummern, 10 Namen, Stichworte oder dergleichen. Die Index-Daten werden im Zuge der weiteren Verarbeitung in Form einer Datenbank weiterverarbeitet und ermöglichen dem späteren Benutzer des Archivs, bestimmte Datensätze wiederzufinden. Die Index-Daten enthalten für alle variablen Datensätze Verweis-Informationen 15 11, welche die Zuordnung bzw. das Wiederfinden der variablen Daten 9 ermöglichen.

Um alle variablen Daten einer Druckseite auch mit dem zugrundeliegenden Formular-Datensatz 8 zu verbinden, muß ein 20 Verweis zwischen variablen Daten und Formular erfolgen. Dieser kann entweder direkt als Verweis 12 oder indirekt über den Index-Datensatz 10 als indirekter Verweis 13 erfolgen. Zur Sicherheit können auch alle drei Verweisarten 11, 12 und 13 gemeinsam für jede Druckseite hinterlegt werden. Die so 25 erzeugte Datenstruktur 7 wird dann über eine geeignete Verbindung dem Archiv-Speicher 3 zugeführt.

Die Formular-Daten 8 und die variablen Daten 9 sind - genauso wie die Archiv-Daten - innerhalb einer Bit Map-Datendomäne 30 14, während die Index-Daten in einem kodierten Format, z.B. ASCII vorliegen.

Figur 2 zeigt eine Ausführungsform für eine erfindungsgemäße Archivierungs-Schnittstelle 1. Sie ist unterteilt in einen Drucker-Controller 6 und einen nachgeschalteten Computer 15, welcher beispielsweise durch einen Personal Computer (PC) 5 sein kann.

Die vom Host-Computer 2 gelieferten Daten werden über einen Kanalanschluß 16 eingespeist. Von diesem werden sie in einem Zwischenspeicher 17 zwischengespeichert. Aus diesem Speicher 10 werden die Daten sukzessive von einer Rastersteuerung 18 entnommen, die die Druckdaten in ein Bit Map Muster umsetzt.

Die Rastersteuerung 18 verfügt hierzu über einen nicht näher gezeigten Speicher, welcher Rasterinformationen zur Umsetzung der Druckdatensprache in die entsprechenden Bit Map 15 Informationen enthält. Die Rastersteuerung 18 unterscheidet innerhalb eines zusammengehörenden Druckjobs Overlay-Daten (Formulare) von variablen Daten. Gleichzeitig produziert sie zu jedem Job einen Index-Datensatz, welcher zur Verwaltung der zu archivierenden Daten dient. Die Rastersteuerung 18 20 speichert die variablen Daten in einem variablen Datenspeicher 19, die Formulardaten in einem Formulardatenspeicher 20 und die Index-Daten in einem Index-Datenspeicher 21 ab. Diese drei Speicher 19, 20 und 21 sind Bestandteil des flüchtigen Arbeits-Bildspeichers 22, des 25 Controllers 6. Die in den Datenspeicher 19, 20 und 21 erzeugten Daten werden dann über eine gemeinsame Schnittstelle 23 an den nachgeschalteten Personal Computer 15 übertragen. Den Datentransfer steuert eine Gesamtsteuerung 24 des Controllers 6, welche mit einer Bedienoberfläche und 30 Steuerung 25 des Personal Computers 15 verbunden ist.

Die Rastersteuerung 18 behandelt Formular-Daten und variable Daten unterschiedlich. Für Formulardaten werden eigene Bit

Maps im Formularspeicher erzeugt. Für die variablen Daten werden im Speicher 19 Bit Maps erzeugt, die keine Formular-Daten (Overlays) enthalten.

5 Die Index-Daten, welche im Index-Speicher 21 abgelegt werden, enthalten Verwaltungsinformationen, welche u.a. für jede Seite des Druckjobs die verwendeten Formular-Overlays ablegt. Dabei wird auch der Bezug zwischen variablen Daten und Formular-Daten hergestellt.

10

Innerhalb des Personal Computers 15 werden die aus dem Speicher für variable Daten 19 und dem Speicher für Formulardaten 20 stammenden Daten in einer Komprimierungsstufe 26 komprimiert. Die Index-Daten werden vorzugsweise ohne Kompression in einer Index-Warteschlange 27 zwischengespeichert, bis sowohl variable als auch Formulardaten die Komprimierung durchlaufen haben. Im Anschluß an die Komprimierungsstufe 26 sind zwei parallel geschaltete Warteschlangen 28, 29 für die variablen Daten und 20 für die Formular-Daten vorgesehen, in denen die entsprechenden Daten so lange zwischengespeichert werden, bis die zugehörigen übrigen Daten die Komprimierungsstufe ebenfalls durchlaufen haben.

25 Die in den Warteschlangen 27, 28 und 29 befindlichen Daten werden dann von einer Synchronisierungsstufe 30 gemeinsam entnommen und über eine Netzwerksteuerung 31 zur Ausgabeschnittstelle 32 in Richtung zum Archivspeicher 3 abgegeben.

Patentansprüche

1. Verfahren zur elektronischen Archivierung des von einem Computer (2) abgegebenen Datenstroms (5), der grafische und/oder Textinformationen enthält, wobei der Druckdatenstrom (5) von einem druckerspezifischen Datenformat in ein auf Pixel basierendes Datenformat (Bit Map) umgesetzt wird, dadurch gekennzeichnet, daß bei den auf Pixel basierenden Daten Formular-Daten (8) von variablen Daten (9) unterschieden werden und diese beiden Datenarten (8, 9) jeweils unterschiedlich verarbeitet werden.
- 15 2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß den variablen Daten (9) Verweise zu den Formulardaten (8) zugeordnet werden.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß ein Formular-Datensatz gleichartiger Formulare einer Gruppe für die Gruppe nur einmal abgespeichert wird, während die zugehörigen variablen Daten aller Datensätze der Gruppe jeweils alle gespeichert werden.
- 25 4. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß die Formulardaten im Archivspeicher (3) nicht abgespeichert werden.
- 30 5. Verfahren nach einem der Ansprüche 1-3, dadurch gekennzeichnet, daß das ursprüngliche Pixelbild aus den Formular-Daten (8) und den variablen Daten (9) rekonstruiert wird.

6. Verfahren nach den Ansprüchen 2 und 5, dadurch gekennzeichnet, daß unter Verwendung der Verweise eine Überlagerung der Formular-Daten (8) und der variablen Daten (9) erfolgt.
- 5
7. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß ein Index-Datensatz erzeugt wird.
- 10 8. Verfahren nach Anspruch 7, dadurch gekennzeichnet, daß der Index-Datensatz einen Verweis auf die variablen Daten (9), und insbesondere auf die Formular-Daten (8) enthält.
- 15 9. Gerät zur elektronischen Archivierung des von einem Computer abgegebenen Datenstroms (5), der grafische und/oder Textinformationen enthält, wobei der Druckdatenstrom (5) von einem druckerspezifischen Datenformat in ein auf Pixel basierendes Datenformat (Bit Map) umgesetzt wird, dadurch gekennzeichnet, daß eine Archivierungs-Schnittstelle (1) vorgesehen ist, die bei den auf Pixel basierenden Daten die Formular-Daten (8) von den variablen Daten (9) unterscheidet und jeweils unterschiedlich verarbeitet.
- 20
- 25 10. Gerät nach Anspruch 8, gekennzeichnet durch einen Drucker-Controller (6), der über eine gemeinsame Schnittstelle variable Daten (9), Formular-Daten (8) und Index-Daten (10) an einen Weiterverarbeitungscomputer (PC) übergibt.

Zusammenfassung

Verfahren und Gerät zur elektronischen Archivierung eines Computer-Datenstroms

5

Die Erfindung betrifft ein Verfahren und ein Gerät zur elektronischen Archivierung des von einem Computer (2) abgegebenen Datenstroms (5), der grafische und/oder Textinformationen enthält. Der Druckdatenstrom (5) wird von einem druckerspezifischen Datenformat in ein auf Pixel basierendes Datenformat (Bit-Map) umgesetzt. Bei den auf Pixel basierenden Daten werden Formular-Daten (8) von variablen Daten (9) unterschieden und diese beiden Datenarten (8, 9) jeweils unterschiedlich verarbeitet.

- 10 15

Figur 1

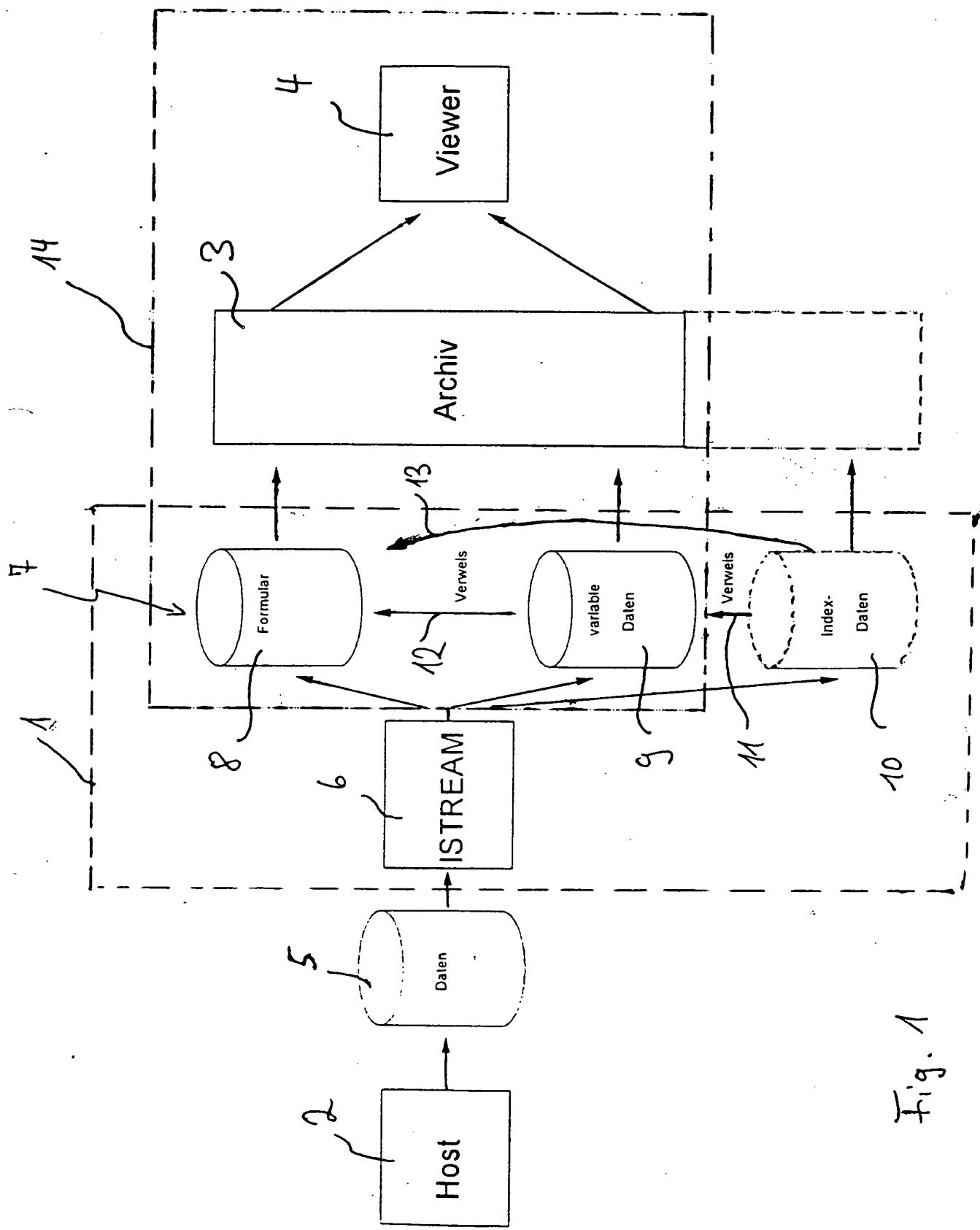


Fig. 1

3080t6

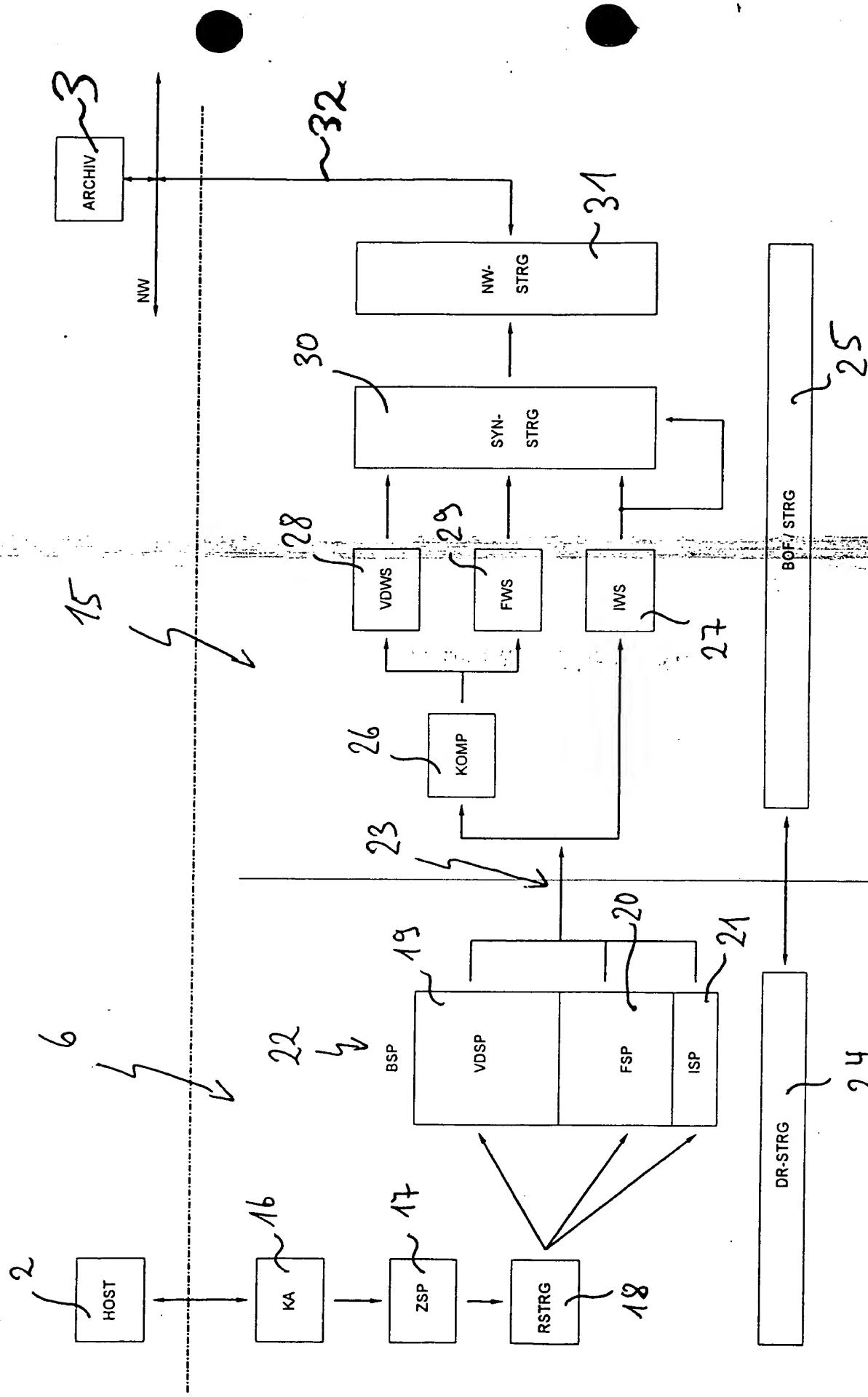


Fig. 2

970803

THIS PAGE BLANK (USPTO)